The Early Years of Panel Data Econometrics

Ariane Dupont-Kieffer and Alain Pirotte

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

> We shall find that two individuals, or the same individual in two different time periods, may be confronted with exactly the same set of specified influencing factors *x* . . . , and still the two individuals . . . may try to remove such discrepancies by introducing more "explaining factors," *x*. But, usually, we shall soon exhaust the number of factors which could be considered as *common* to all individuals, and which, at the same time, were not merely of negligible influence upon *y*. The discrepancies $y - y^*$ for each individual may depend upon a great variety of factors, these factors may be different from one individual to another, and they may vary with time for each individual.
>
> —Trygve Haavelmo, "The Probability Approach in Econometrics" (1944)

Over the last fifty years, a major shift in applied econometrics has been the increasing use of panel data. A panel data set is usually defined as repeated observations on the same economic units (individuals, firms, households, countries, states) over several time periods. Panel data may be contrasted with cross-section data, which contain observations on individual units at a point in time, and with time-series data, which contain observations, usually of an aggregate nature, over time without any individual dimension. A well-known example of a panel data set is the Panel Study of Income Dynamics (PSID) collected by the Institute for Social

We would like to thank two anonymous referees, Patrick Sevestre, Marcel Boumans, John Aldrich, and the participants of the *HOPE* 2010 conference for their useful comments and suggestions.

History of Political Economy 43 (annual suppl.) DOI 10.1215/00182702-1158754
Copyright 2011 by Duke University Press

Research at the University of Michigan.[1] The PSID provides mostly economic and demographic data.

As just noted, panel data econometrics has been one of the most productive parts of the field. The increase in publications related to panel data gives an idea of this success. According to the Social Sciences Citation Index, there were 70 articles related to panel data in 1990, 375 in 1995, 716 in 2000, 1,165 in 2005, and 1,788 in 2008. Thus the number of articles published increased by more than twenty-five-fold between 1990 and 2008. These developments were gathered in the third edition of the well-known and widely referred to *Panel Data Handbook* edited by László Mátyás and Patrick Sevestre and published in 2008.[2]

The attractiveness of panel data is rooted in the expectation for more "realistic" dynamic specifications and better understanding of microeconomic behaviors. Two founding fathers of panel data econometrics reminded readers of this point years after their seminal article was published in *Econometrica* in 1966: "One of the main reasons for being interested in panel data is the unique possibility of uncovering disaggregate dynamic relationships using such data sets" (Balestra and Nerlove 1992, 16). One of them emphasized that "not only do panel data frequently provide the opportunity for introducing many more explanatory variables and more complicated dynamics, but they also permit us to model more explicitly the latent disturbances themselves as components common to all individuals at a point in time and as time-persistent components" (Nerlove 2002, 6).

Moreover, the use of microdata for macroeconomic purposes is nowadays a field of important applied economic research. Since the pioneering papers by Yair Mundlak (1961) and Pietro Balestra and Marc Nerlove (1966), panel data econometrics has produced general specifications and methods that have become basic tools for applied researchers. These methods have helped make panel data econometrics a specific branch of econometrics. For over forty years, the concepts of fixed and random effects have been essential for linear and nonlinear models that consider more than one dimension.

This interest in panel data is partly related to the developments in economic theory, the developments in computer technology and software

1. The PSID began in 1968 with forty-eight hundred families and grew to more than eight thousand families in 2005. By 2005 the PSID had collected information on more than seventy thousand individuals, spanning as much as thirty-seven years of their lives.

2. First edition published in 1992.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

programs, the progress in the elaboration and implementation of appropriate statistical and econometric methods, and the availability of panel data sets. But beyond this conjunction of elements and favorable context, what can explain the development of panel econometrics per se and its specificities?

At the beginning of this history of panel econometrics, much attention seems to have been devoted to formulating econometric models relevant to important economic issues and data rather than pure methodological issues. Irving Hoch (1962, 34) summarized the way to proceed:[3]

> The theoretical development involves three stages: (a) construction of an equation set—or economic model—that describes the behaviour of the competitive firm; (b) derivation of a statistical model from the economic model by the introduction of disturbance terms and by the specification of characteristics of those disturbances; (c) further development of the statistical model, calling for the use of combined time-series and cross-section data.

This focus on economic modeling represents a break with the current practice of econometricians under the Cowles Commission paradigm. The early econometricians driven by Ragnar Frisch, before World War II, devoted much of their time to identification issues and tried to deal with an economic approach to the unobservable phenomena. The analysis of the error was central in their "econometric" investigation of the latent variables being understood as the "unification of economic theory, mathematics and statistics" (Frisch 1933, 1). But the thesis of Trygve Haavelmo published in 1944 allowed or was understood by the community of econometricians, then gathered under the Cowles Commission banner. By then, the treatment of errors was mainly statistical and methodological. Assumptions being made were more on the statistical properties than on their economic content. The specificity—and the break in the history of econometrics—we identify for panel econometrics is precisely a renewed interest in the error term as a potential source of information of economic phenomena and a potential help to formulate economic laws.

As emphasized by Jeff Biddle (this volume), agricultural economics offered resources and tools to investigate this identification issue. Indeed, one of the key forerunners, Nerlove (2002), emphasized that the first panel

3. Hoch wrote his PhD dissertation, "Estimation of Agricultural Resource Productivities Combining Time Series and Cross–Section Data," at the University of Chicago in 1957. A progress report was presented at the Econometric Society meeting in Montreal in September 1954 and reported in Hoch 1955.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

econometricians turned to already existing model principles used in other fields to solve their specific problems related to panel data.[4] For example, Hoch (1962) estimated a Cobb-Douglas production function relating the output value in dollars to the value of inputs in four categories of inputs using a panel of sixty-three Minnesota farms observed for 1946–51. In this perspective, he used the analysis of covariance incorporating time and individual fixed effects. But Clifford G. Hildreth (1949, 1950) and Hoch (1962) were still embedded in the dominant Cowles Commission approach. Even if they imported statistical approaches borrowed from Ronald Fisher's corpus and considered no longer the error term as a "nuisance," they were still mainly driven by a methodological preoccupation rather than an economic investigation. However, Hoch raised the question at the theoretical level of the nature of the latent variables.

Following this path opened by Hoch,[5] Mundlak (1961) also worked on estimating farm production function. An important methodological and theoretical question behind these first applied papers was the problem of latent variables. Thus, on panel data, the crucial aspect of the problem is to get a clear understanding of how differences in behavior across individuals and/or through time could and should be modeled. In this perspective, the first econometricians have tried to control the heterogeneity of individual behaviors related to unobservable components. The investigation of heterogeneity was allowed by the panel data set itself. Indeed, these sets provide mainly more informative data, variability, and efficiency.

The present essay focuses on the early years of panel data econometrics. Section 1 presents the context before the emergence of panel data econometrics through the pioneering papers by Mundlak (1961) and Balestra and Nerlove (1966), while section 2 analyzes the question of specification errors and covariance analysis in the light of Mundlak 1961. Section 3 focuses on dynamics and individual heterogeneity following Balestra and Nerlove 1966, and the last section concludes.

1. Identification and Estimation: Early Attempts

Panel econometrics emerged with specific concerns about the treatment of data and about the nature of the economic (and not statistical) relations

4. The fixed- and random-effects models have a long history in astronomy, agronomy, and statistics, going back to the nineteenth century.

5. Mundlak mentioned Hoch 1955 in a footnote on the first page of his article.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

among variables, one of those concerns being, for example, the investigation of the economic process that gives rise to serial correlation. These concerns were much closer to the ones that justified the research program of the early econometricians at the beginning of the twentieth century.

In the late twenties and early thirties, one main topic of interest to the first econometricians was estimation of demand and supply curves. The difficulties encountered drove them to think about defining relations or more precisely correspondence rules between economic phenomena, available data, and theoretical frameworks. Debates arose on the nature of data and the role of observation in defining economic theories.

In that context identification and estimation procedures were mixed and not clearly distinguished. Identification is usually divided into four tasks: model definition, identification of variables, estimation, and testing. As shown by Mary Morgan (1990, chap. 6), early econometricians did not separate these tasks. Their main concern was how to deal with data and connect economic theory and economic life. This preoccupation was asserted in the editorial of the first issue of *Econometrica*: "Theory, in formulating its abstract quantitative notions, must be inspired to a larger extent by the technique of observation. And fresh statistical and other factual studies must be the healthy element of disturbance that constantly threatens and disquiets the theorist and prevents him from coming to rest on some inherited, obsolete set of assumptions" (Frisch 1933, 2).

How to go forward from observation to explanation? Modeling and confluence analysis were the answers proposed by the founder Frisch and later by Haavelmo.[6] Modeling work was driven by economic theory and the need to identify economic laws based on economic data investigation and analysis. The econometric agenda was based on the identification of constant terms among the observations and the explanation of the gap between observed data and constant or autonomous data. Frisch urged, and later Haavelmo, for the definition of a specific identification methodology relevant for economics. As emphasized by Morgan (1990, 189), on the basis of Haavelmo 1943 and 1944, econometricians moved away from considering the definition of correspondence between mathematical economics and statistical economics:

> After this work of the late 1920s and early 1930s, the theoretical problem of identification was not taken up again until the work of Koopmans and others at the Cowles Commission in the 1940s. This work,

6. See Morgan 1990, Qin 1989, and Epstein 1987.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

> stimulated by Frisch's (1938) paper on autonomous relationships and business cycles [Frisch (1938) 1995], led to the codification of the rank and order conditions for identification of linear models involving several equations (see Qin (1989) and Epstein (1987)). Their other advance was in dealing with the problem of overidentification, whereas the work of the 1920s and 1930s had dealt with the cases of just-identified and underidentified models (see Koopmans (1949 and 1950)). This codification and formalisation of the identification problem transformed it into a technical problem divorced from the other correspondence problems of location, interpretation and even model choice, of which in the 1920s and 1930s it was seen to be a part.

Econometricians afterward favored more technical questions on the treatment of errors and on estimation methods and issues. Thus, in the nineteenth and early twentieth centuries, the interpretation of the error term became for the econometricians a central issue, and they turned to debate on variance and covariances raised in agriculture, astronomy, and statistics. With the adoption of Haavelmo's methodology by the Cowles Commission, econometricians moved away from the errors-in-variables (Morgan 1990) perspective to adopt an errors-in-equation perspective. Duo Qin and Christopher Gilbert (2001) showed that this tendency drove econometricians to develop simultaneous equations. The Cowles Commission methods for estimating structural equations allow assuming that the error terms of the structural equations were independently distributed and that their causes were too small or negligible. They mainly focused on static analysis and considered the dynamic pattern in time-series analysis a "nuisance." More specifically they showed that

> His distinction [Frisch's distinction in Frisch (1938) 1995 between stimuli (structural shocks) and aberrations (nonstructural disturbances)] was lost in the Cowles Commission work. Subsequently, the general perception of errors regressed to the initial view, i.e., that they lack economic significance. This approach was strengthened by the assumption, made for statistical convenience, that the errors followed the serially independent and identical distribution (i.i.d.), on the argument that these errors merely represented the aggregate effects of a large number of individually unimportant omitted variables. (Qin and Gilbert 2001, 425)

The Cowles Commission did not undertake the interpretation of errors as random shocks and focused on developing statistical tools to identify and deal with possible autocorrelation among the residuals.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

In the fifties and sixties, those who later became known as panel econometricians brought the debate of identification and estimation back to investigating the economic meaning of latent variables and to focusing more on the identification issues than on the technical aspects of estimation procedures. They put forward two elements—heterogeneity and dynamics—for consideration to explain the gap between the structure and the observed data and to investigate latent variables. They were then compelled to focus on analysis of variance and covariance and turned to other scientific fields (such as agriculture and astronomy) that had already faced heterogeneity issues.

The solution was to distinguish between fixed-effects models (first defined by Fisher in 1925) and random-effects models (traceable to Airy 1861 and Chauvenet 1863). Henry Ellis Daniels (1939) clearly established the difference between the two approaches; but the importance of that distinction for experimental versus nonexperimental data was put forward for applied research by Churchill Eisenhart only in 1947.

Daniels (1939) and Eisenhart (1947) took over the debate on the nature of the error term and on variance analysis. More broadly, the issue was how to deal with latent variables. The break with the Cowles Commission paradigm happened when some econometricians applied this distinction with Hildreth 1950, Hoch in the midfifties, and then with Mundlak 1961, and Balestra and Nerlove 1966. Nerlove (2002, 4–18) acknowledged his debt to statistical development in variance and covariance analysis. These authors are the heirs of Fisher's (1925a) contributions, but they clarified the debate on the use of fixed- versus random-effects models according to the nature of the data. They imported this approach but made it their own by using it not only as a statistical improvement to econometric methodology. The difference between Hildreth/Hoch and Mundlak/Balestra-Nerlove is the purpose behind the use of the fixed- and random-effects models. Indeed, this reference to Fisher's methodology was clearly for all a way to reintroduce concerns on unobservable variables. But in the case of the first ones, the concern was still methodological; in the case of the latter, the concern was more on the economic nature of the latent variables and of the relationships among them. Mundlak and Balestra-Nerlove sought in the error term information about economic phenomena, in particular on the individual level of heterogeneity.

The contributions by Eisenhart (1947) and Henry Scheffé (1956, 1959) were mentioned by the pioneers of panel econometrics and thus constitute fundamental developments for the emergence of panel data econometrics.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

Eisenhart (1947) started his introduction to a special issue of *Biometrics* on variance analysis by acknowledging the contribution of Ronald Fisher to analyzing idiosyncrasy among specific phenomena as agricultural or biological. He then questioned the relevancy of the use of this particular methodology for variance analysis in other scientific fields:

> The statistical technique known as "variance analysis" developed more than two decades ago by R. A. Fisher to facilitate the analysis and interpretation of data from field trials and laboratory experiments in agricultural and biological research, today constitutes one of the principal research tools of the biological scientist, and its use is spreading rapidly in social sciences, the physical sciences and in engineering. (Eisenhart 1947, 1)

His aim was not only pedagogical by offering a step-by-step description of how to proceed; it was also explanatory, as he explained to which data this variance analysis can be applied:

> The principal deficiency of these books [handbooks in statistics] has been their failure to state explicitly the several assumptions underlying the analysis of variance, and to indicate the importance of each from the practical point of view. . . . My assignment is to enumerate the several assumptions underlying the analysis of variance and to point out the practical importance of each. (2–3)

The "assumptions" were related to the nature of the phenomena at stake, and the analysis of variance will differ from one category to another:

> Turning now to my assignment, I am obliged at the outset to draw attention to the fact that analysis of variance can be, and is, used to provide solutions to problems of two fundamentally different types. These two distinct classes of problems:
>
> —Class I: Detection and Estimation of Fixed (Constant) Relations Among the Means of Sub-Sets of the Universe of Objects Concerned. This class includes all the usual problems of estimating, and testing to determine whether to infer the existence of, true differences among "treatment" means, among "variety" means, and, under certain conditions, among "place" means . . . the analysis-of-variance are the least-squares solutions. The cardinal contribution of analysis of variance to the actual procedure is the analysis-of-variance table devised by R. A. Fisher . . .

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

> —Class II: Detection and Estimation of Components of (Random) Variation Associated with a Composite Population. This class includes all problems of estimating, and testing to determine whether to infer the existence of components of variance ascribable to random deviation of the characteristics of individuals of a particular generic type from the mean values of these characteristics in the "population" of all individuals of that generic type, etc. In a sense, this is the true analysis of variance . . . . Problems of this class received considerably less attention in the literature of analysis of variance than have problems of Class I. (3–4)

One main issue that econometrics had to face was precisely the Eisenhart Class II problems. Eisenhart took over the debate opened after Fisher's publication in 1925 on the treatment of experimental data versus nonexperimental data. In some ways, he indirectly validated Frisch's position toward Fisher's analysis of variance.

Indeed, econometrics had to develop its own methodology. According to Frisch's point of view, the confluence analysis was able to help econometricians facing nonexperimental data when Fisher's analysis of variance was relevant for experimental data. Frisch focused on multiple correlation coefficients as key in determining the presence of perturbations and in distinguishing them from accidental variations. From the beginning of his scientific career in 1925 but more clearly after 1929, Frisch pointed out the need for specific statistical methodology to deal with multicollinearity, heterogeneity, and dynamics in explaining residuals. Writing this history, Eisenhart (1947, 19) appears to be part of this tradition and of the definition of a specific methodology for nonexperimental data: "The answer [whether the parameters of interest specify fixed relations or components of random variation] depends in part, however, upon how the observations were obtained; on the extent to which the experimental procedure employed sampled the respective variables at random. This generally provides the clue."

In the econometric literature, the idea of distinguishing between fixed and random effects was taken over by Hildreth (1950). He related unobserved characteristics to individual effects combined or not with time effects: "He set out a three-component model for the latent disturbances in a simultaneous-equations model and considered estimation when these components might be considered random or when two of them, period effects and individual effects, might be considered fixed effects and thus parameters to be estimated" (Nerlove 2002, 17).

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

Hildreth's contribution should be seen as an early attempt to identify heterogeneity among individuals and periods, including constant vectors as fixed variables to enlighten the presence of "latent" individual effects and then to introduce and to combine fixed-effect and random-effect modeling. His purpose was clearly to tackle with heterogeneity mainly the possible omitted variables, that is, issues to look at when dealing with unobserved phenomena: "It may be believed that there are unobserved individual characteristics which cause individuals to act differently and which are persistent over time. There may be observed influences that affect individuals in pretty much the same way but change over time" (Hildreth 1950, 2).

First, he thought about combining time-series data and cross-section data and tried to define the analysis of variance in that case. Working out a better formulation of the error term, he appeared to be dissatisfied with the current analysis of variance and more broadly with how to formulate it within structural equations: "I find it difficult to choose between the alternatives of allowing for these variations peculiar to individuals and variations peculiar to time though fixed parameters or through random parameters" (2). He proposed a new way to integrate these two aspects of differences, individual and time patterns, and emphasized that the maximum likelihood was difficult to derive.

As Hildreth noted in his introduction, there were two innovative aspects of the approach developed then by these pioneers:

> Two sorts of contributions to the problem of estimating economic relations from empirical data may be expected from the joint use of cross-section data and time series. The investigator can expect to work with large numbers of observations thus reducing his sampling errors and making tests of significance more powerful. He can also choose from a wider selection of statistical models thus having a better chance to construct a model that is both realistic and manageable. (1)

This paragraph is of interest to our history for two reasons. First, it shows that one of the innovations and fruitful perspectives was also the collection of data and the ability to get a sample of a larger size to control heterogeneity and dynamic patterns. Second, it shows that Hildreth and Hoch were still concerned by statistical modeling and not economic modeling to deal with heterogeneity issues and identification challenges, which was the case with Mundlak and Balestra-Nerlove.

Indeed, the question of heterogeneity was already raised by Jacob Marschak and William H. Andrews (1944) on estimating production function.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

They put forward heterogeneity of the inputs as a factor explaining differences between firms. They dealt directly with identification issues raised by Haavelmo (1944) (Marschak and Andrews 1944, 147, 151) and focused on how to incorporate or more precisely capture the variations in production functions among firms owing to time effects and the intrinsic specificities of each firm (145). They set up their paper as an answer to the fact that there were no experimental data in economics (143). They sought to isolate differences owing to the firm (technology) or to time and advocated that the variance analysis could lead to a way to isolate homogeneous groups within the data to compare them (173–74).

Besides this heterogeneity issue, their paper proposes a solution to aggregation bias. Indeed, panel data allow circumventing problems owing to aggregation for estimating parameters defined at the individual level (firms or agents). The authors even argued that knowledge of the distribution of the heterogeneity might solve the identification issue related to estimating simultaneous equations with aggregated times series.[7]

Later, Hildreth (1949, 1950) showed the path to be taken and brought back identification issues to the forefront of the debate. Hoch (1955, 1957, 1958, 1962) attempted to address the estimation and identification issues raised by Marschak and Andrews in 1944 by combining time-series and cross-section data to capture the heterogeneity among economic determinants. He developed what is now known as the analysis of covariance. Hoch concludes that “management” can be one of the left-out factors. Beyond the statistical problem, Hoch brought back the debate on residuals and variance and covariance to theoretical questions on the nature of latent variables. We see here that he was moving from the estimation issues to identification issues by trying to find theoretical explanations to the bias (and a way to identify the bias). This is the beginning of numerous papers founded on the same idea, especially the seminal article by Mundlak (1961).

2. Specification Errors and Covariance Analysis

At the early stages of panel data econometrics much attention seems to be devoted to formulating econometric models relevant to important economic issues. The first papers using panel data were empirical studies on

7. See especially Marschak and Andrews 1944, 151, and sec. 3, Data and Findings, sec. 24, 169–70.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

firm behavior. Mundlak's (1961) paper was one of these. More specifically, Mundlak's questioning was essentially to evaluate the contribution of the production factors, given the (unobservable) heterogeneity that characterizes firms. In this perspective, he faced a major problem: "It has been felt for a long time that the estimates of the parameters of production functions are subject to bias as a result of excluding the variable which represents management" (Mundlak 1961, 44). The problem of bias was already emphasized by Zvi Griliches (1957, 13) when he said: "The specification error conceded most often by estimators of production functions is the omission of entrepreneurial or managerial services."

Griliches (1957) showed that the omission of managerial inputs from the Cobb-Douglas production function biases the elasticity of output with respect to capital inputs and the estimate of returns to scale. Mundlak (1961, 44) was faced with two kinds of problems: the heterogeneity of entrepreneurial capacities and the "lack of units for its direct measurement." He assumed that "whatever management is, it does not change considerably over time; and for short periods, say a few years, it can be assumed to remain constant."

Mundlak dealt with the log-linear form of a Cobb-Douglas production function including a variable "management" that varies over individual i but constant over time t (M_i). He retained a fixed individual-effects specification:

$$y_{it} = b_0 + b_1 x_{1it} + \cdots + b_k x_{kit} + \mu_i + \varepsilon_{it} \quad i = 1, \ldots, N, t = 1, \ldots, T, \qquad (1)$$

where y_{it} was the dependent variable (logarithm of output), x_{jit} was the explanatory variable j (logarithm of input j), $j = 1, \ldots, k$, $b_0, b_1, \ldots, b_k$ were the coefficients of explanatory variables to be estimated, ε_{it} was the disturbance, $\mu_i = cM_i$ with

$$\sum_{i=1}^{N} \mu_i = 0. \qquad (2)$$

He had no observations to characterize M_i. Nevertheless, the "management" could be associated to an arbitrary constant measure, namely, μ_i. So the fixed individual effects were proportional to a latent variable measuring "management." Mundlak argued that, along the lines suggested by Hoch (1955), a panel of firms for which the "management" factor could be assumed to be approximately fixed over time for each firm can be used to obtain unbiased estimates of the "intrafirm" production function. Under

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

those assumptions, the statistical approach chosen by Mundlak was the analysis of covariance to get unbiased estimates of the elasticities of the Cobb-Douglas production function.[8] Estimates of individual effects μ_i, $i = 1, \dots, N$, were obtained from the relation

$$\hat{\mu}_i = (y_{i.} - y_{..}) - \sum_{j=1}^{k} \hat{b}_j (x_{ji.} - x_{ji..}) \tag{3}$$

with

$$y_{i.} = \frac{1}{T}\sum_{t=1}^{T} y_{it},\ y_{..} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it},\ x_{ji.} = \frac{1}{T}\sum_{t=1}^{T} x_{jit},\ x_{j..} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{jit}, \tag{4}$$

and $\hat{b}_j$ was the "intrafirm" estimator. One main advantage of this approach was that it allowed an economic interpretation of the fixed individual effects. The estimated coefficients were obtained from a panel of sixty-six farms in Israel, for 1954–58 (i.e., a short time period of five years). The elasticities, $\hat{b}_j$, $j = 1, \dots, k$, were based on the variations "within" firms (Mundlak 1961, 47). "If the assumptions of classical regression hold and if the function is completely specified, then the estimates obtained are unbiased and best" (47).

Moreover, Mundlak identified the parameter c to derive an estimate of the "management" variable under the assumption of the divisibility of the factors. Also, he pointed out that the individual fixed effects (μ_i) could be a mixture of several components like a farm effect, which did not depend on "management." In this case, a specific parameter is added to account for farm effect. He also suggested introducing fixed-year effects to take into account variations in the level of productivity that occurred in time. Finally, he presented four sets of estimates of the Cobb-Douglas function obtained under the following assumptions:

neither a fixed time nor a fixed firm effect;
a fixed time effect but without fixed individual effect;
no fixed time effect but with fixed individual effect;
the unrestricted regression allowing for both a fixed time and individual effects.

He compared the fixed-time and individual-effects regression with other regressions that did not include simultaneously those two effects. The gap

8. Mundlak cited Scheffé 1959 as his statistical authority.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

between estimation results characterized the "management bias" (Mundlak 1961, 51). He compared the estimates of pooled and fixed individual-effects models, and those of fixed-time effects and fixed-individual and time-effects models. For each case, he computed the absolute and relative bias: "In both sets of comparison the firm effect turns out to be highly significant. The implication is that the usual regression which is computed by not allowing for the firm effect is likely to be subject to bias. The rejection of the hypothesis of no firm effects is a necessary condition of management bias" (51).

These results went beyond those of Hoch (1955, 326) who noted that a possible explanation for difference in estimates was the omission of "entrepreneurial capacity." Thus, in this first paper, Mundlak focused his comments mainly on specification errors and identification, not strictly on statistical aspects.

3. Dynamics and Individual Heterogeneity

The title of the seminal paper by Balestra and Nerlove (1966, 585) contains two main dimensions: "Pooling cross section and time series data in the estimation of a dynamic model: the demand for natural gas." They pointed out that "the more specific problem of estimating the parameters of demand function, when the demand is cast in dynamic terms and when observations are drawn from a time series of cross sections. Accordingly, this paper is centered around these two major themes, although, as the title suggests, the emphasis is placed on the second one"; over thirty years later Nerlove (2002, 27) emphasized:

> Early on in the development of panel data econometrics, it was widely recognized that dynamic panel models are of key importance and, indeed, it is the need to estimate dynamic models that differentiates the econometric problems from those generally discussed in the general statistical literature on variance components and on the analysis of covariance. Elsewhere, Balestra and I (Nerlove and Balestra, 1996) have argued that all models of economic behavior are basically dynamic, whether or not the dynamics is explicit or not. The appropriate dynamic may be modeled by formulating the relation to be estimated as a difference equation or by modelling stocks and flows explicitly with multiple-equation models or by admitting the possibility that the errors may themselves be correlated over time or both.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

To introduce a dynamic specification, Balestra and Nerlove (1966) investigated the demand for natural gas. Nevertheless, the essence of their approach "is not restricted to the gas model" (585). At the beginning of the paper, the main idea consisted in formulating a demand function of natural gas where consumption was connected to the stock of gas appliances. So this demand should incorporate a stock effect and some assumptions about the adjustment of these stocks over time. An appropriate approach for the authors to take into account these aspects was a dynamic model. So, the paper's first section was focused on dynamic mechanism, not on pooling cross-section and time-series data (hereafter panel data). Under several assumptions, an autoregressive model was derived:[9]

$$y_{it} = b_0 + \rho y_{i,t-1} + b_1 x_{1it} + \cdots + b_k x_{kit} + u_{it}. \tag{5}$$

The type of data appeared only on page 589 when they mentioned that "the investigation reported here is based on data by state and covers the period 1950–1962." The model was estimated from a panel of thirty-six U.S. states for 1950–62 (i.e., thirteen years): "All observations are grouped together and estimation . . . is performed on the combined sample of cross section and time series" (Balestra and Nerlove 1966, 589).

Nevertheless, first estimates were made for 1957–62 (mature stage of gas market) under the assumption of homogeneity of the U.S. states: "During the later period (1957–62), all states included in the sample are reasonably homogeneous as far as gas availability is concerned, and the assumption of perfectly elastic supply is approximated" (590).

They were not satisfied with the first results, especially the coefficient of the lagged dependent variable, which implied an implausibly negative depreciation rate of gas appliances. Several explanations were presented. One of them concerned the consequences of unobservable heterogeneity: "One possible explanation for the results thus far obtained is that, when cross section and time series data are combined in the estimation of a regression equation, certain 'other effects' may be present in the data. A natural way to account for these 'other effects' is to introduce explicitly into the equation individual shift variables" (592).

9. This autoregressive form is associated with the gas demand obtained by Balestra and Nerlove (1966, 589). We can also consider a more general case where the autoregressive part is of order p (i.e., p lagged values on the dependent variable), as in Balestra and Nerlove 1966, 594.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

By pooling cross-section and time-series data, it seems that Balestra and Nerlove were aware of two potentials of this "huge" sample size: it allows not only controlling mainly individual heterogeneity but also treating explicitly the individual pattern in the model. So they put fixed individual effects in (5), but the results were not economically plausible:

> The rationale of this procedure is that the data contain an additive effect specific to the individual (state). To account for such effects, dummy variables corresponding to the 36 different states may be introduced explicitly into the model. It is moot, however, whether the dummy variable method is appropriate in the case of a dynamic model. The presence of lagged endogenous variables may make it difficult, if not impossible, to separate the individual (state) effects from the effect induced by the lagged variable. (592)

More precisely, the estimated value of the lagged endogenous variable implied a rapid depreciation rate of gas appliances of over 30 percent, which is highly implausible. The authors explained that the lagged endogenous variable partly reflected a regional effect rather than a true lag effect. Moreover, they advanced that the fixed effects wasted and contributed to obtain a reduced value of the coefficient of lagged endogenous variable.[10] Their idea was to introduce individual effects in an alternative fashion to overcome this difficulty. Nerlove (2002, 22) said years later: "I recall rather pedantically explaining that the disturbances represented numerous, individually insignificant variables affecting the gas consumption in a particular state in a particular year, some of which were peculiar to the state (i.e., state-specific), and didn't change much or at all over time."

They proceeded to formulate what we call now a two-way error components model already introduced by Hildreth (1950), namely, "residual model" in the article. More exactly, they retained an additive decomposition of the disturbance:

$$u_{it} = \mu_i + \lambda_t + \varepsilon_{it}. \tag{6}$$

Nevertheless, the authors considered finally only an individual effect, not a time effect. They mentioned that "this would greatly complicate the analysis without adding any essential generality" (Balestra and Nerlove 1966, 594). Section 3 of the paper described the implications of the specific structure (6) of the disturbance. This structure led to the usual

10. Nickell (1981) understood and demonstrated why such results were obtained.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

block-diagonal residual variance-covariance matrix, dependent on two unknown parameters σ_μ^2 and σ_ε^2. In matrix form we have

$$E[uu'] = \sigma_u^2 \begin{pmatrix} A & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A \end{pmatrix} = \sigma_u^2 \Omega, \tag{7}$$

$$A = \begin{pmatrix} 1 & \eta & \cdots & \eta \\ \eta & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \eta \\ \eta & \cdots & \eta & 1 \end{pmatrix}, \eta = \frac{\sigma_\mu^2}{\sigma_u^2} \text{ and } \sigma_u^2 = \sigma_\mu^2 + \sigma_\varepsilon^2. \tag{8}$$

A is an equicorrelated matrix where η was called the intraclass correlation by Fisher (1918, 1925a, 1925b). A great deal of Balestra's 1965 dissertation,[11] and Balestra and Nerlove 1966, was to present various alternative methods of estimation of η, which was unknown. Moreover, over thirty years later Nerlove (2002, 23) emphasized the problem at that time to compute the inverse of Ω:[12]

> Inverting such large matrix, even were [η] known, would have been a problem for us at that time. Wallace and Hussain (1969) and Henderson (1971) were yet to come. However, the matrix Ω has a rather simple structure despite its large size. Balestra was then, as he has ever been, a wiz with matrices; it took him about a week to find the characteristic roots of Ω and the orthogonal transformation which would reduce Ω to diagonal form.

Balestra and Nerlove (1966, 593) also emphasized "the presence of such lagged values which produces the essential difficulty of the problem, and which distinguishes it from the type of problem discussed in recent econometric literature."

11. His dissertation was published in 1967.

12. In their panel data set, $N = 36$ and $T = 6$, which induced a matrix of dimension ($NT \times NT$) = (216 × 216) for Ω. This kind of transformation was considered by Max Halperin (1951) in the context of no lagged value of the dependent variable in the list of explanatory variables.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

From (5), if $\rho = 0$, no lagged value of the dependent variable was included in the explanatory part. The ordinary least squares (OLS) were unbiased and consistent for estimating the coefficients under usual assumptions. Nevertheless, under (6) the variance-covariance matrix of the coefficients is biased and inconsistent. Balestra and Nerlove (1966, 596) referred to a two-stage procedure proposed by Arnold Zellner (1962) and later by Lester G. Telser (1964) to solve the problem. Following Zellner 1962, in a first step, an estimate of the variance-covariance matrix of residuals was obtained using OLS estimates of (5). In a second step, this matrix was used to derive new estimates of the parameters. They mentioned some characteristics of this two-stage procedure:

> Zellner discusses the asymptotic properties of such two-stage estimators and shows that the gain in efficiency depends on the values of the off-diagonal elements in Ω (in our case, the extent to which [η] differs from zero), and on the correlation of the independent variables for the different individuals (blocks). If the independent variables for each individual are perfectly correlated, Zellner's results show that if no shift variables are included, the asymptotic efficiency of the ordinary least squares estimators is the same as that of the proposed two-stage estimators. Such perfect correlation will rarely be the case, however. (Balestra and Nerlove 1966, 597)

Unfortunately, the other case, that is, $\rho \neq 0$, is not as simple: "When lagged endogenous variables are included among the explanatory variables of y in (5), it is no longer true that the ordinary least squares estimates of [the parameters] are consistent unless there is no serial correlation of any kind. In this case, of course, there is no possibility either of increasing the efficiency of the estimates by any sort of iteration" (597).

They dealt with various alternatives methods of estimation, mainly maximum likelihood (ML) and instrumental variables (IV) approaches. Nerlove (2002, 24) explained why they have preferred the ML estimator:

> I recall that in late 1963 we headed straight for maximum likelihood as the preferred method for estimating [η] simultaneously with the other parameters. It was only because this method seemed to fail that we turned to other alternatives. At the time, however, we didn't realize, as Bhargava and Sargan (1983) were to show us twenty years later, that the presence of a lagged value of the dependent variable as one of the explanatory variables, i.e., the autoregressive nature of the relationship

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

> to be estimated from the panel data, makes all the difference in the formulation of the variable as predetermined, that is, as fixed just like one of the *x*s.

Also, Nerlove recognized that "Hildreth (1949) would have told us this, but we were unaware of this paper at the time" (24).[13] So Balestra and Nerlove (1966) looked for another estimator asymptotically equivalent to the ML estimates. They indicated the inconsistency of OLS to estimate (5): "The reason that ordinary least squares estimates are inconsistent when lagged variables are included is that the variables are correlated with the current values of the residuals u_{it} since they are determined to the same degree as the current value of the dependent variables by μ_i" (Balestra and Nerlove 1966, 603).

These authors also emphasized that the "same sort of difficulty arises in the estimation of one of a system of structural equations involving more than one endogenous variables of the system" (603). In this context, the IV could be an interesting alternative:

> One solution to this difficulty is to use as instrumental variables a sufficient number of other exogenous or (in the absence of serially correlated residuals) lagged endogenous variables appearing elsewhere in the system in the formation of the "normal" equations so that the current endogenous variables in the equation need not be used for this purpose. The difficulty, of course, is that there are usually more than enough predetermined variables for this purpose, and a choice must be made among them. One of Theil's contributions in the development of two-stage least squares was to show how such a choice could be avoided by selecting as instrumental variables those linear combinations of all predetermined variables most highly correlated with the current endogenous variables whose values they replaced in forming normal equations. (604)

So the main question was to find some additional exogenous variables and how they should be obtained: "The key to the solution is to be found in the idea that the lagged values of the dependent variables are determined in a sense by other equations, although these are just lagged versions of the equation we are trying to estimate" (604).

Under certain restrictions, they showed that (5) had the solution

13. It is an interesting point that Nerlove in 2002 acknowledges Hildreth 1949 in retrospect, although Hildreth was not discussing the bias problem created by the use of lagged dependent variable.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

$$y_{it} = b_0^{\bullet} + \sum_{j=1}^{k} b_j \sum_{\tau=0}^{\infty} \rho^{\tau} x_{jit-\tau} + u_{it}^{\bullet}. \qquad (9)$$

From (9), endogenous variable y_{it} appeared as a combination of exogenous variables (i.e., they were independent of the current and past values of the disturbances u_{it}), suggesting that the lagged values of the explanatory variables ($x_{jit-\tau}$) may be used as IV in to generate normal equations for estimating parameters in (5). Finally, in a first step, using IV they got a consistent estimate of η, and in a second step, they used this estimator to obtain plausible estimates of the coefficients in (5): "Application of the residual model to the gas data produces estimates that are in agreement with theoretical expectations. . . . This result lends support to the basic hypothesis embodied in the dynamic model of gas demand" (606).

In this article, the unobservable individual heterogeneity appeared to be a central problem in the context of an autoregressive model. In particular, this implied that the coefficient of lagged dependent variable was biased and inconsistent using OLS (when $N \to \infty$ and T is small). This also induced specific problems when ML is used.

4. Conclusion

The present essay has shown how the first panel data econometricians used already existing model principles to solve their specific problems and imported these principles into econometric analysis. Thus Mundlak (1961) evaluated the contribution of the production factors, including entrepreneurial capacities using a fixed-effects model. Later, in a dynamic framework, Balestra and Nerlove (1966) retained a random-effects model as an alternative to fixed-effects specification. Both specifications captured unobservable individual heterogeneity, which was an important characteristic on panel data. They tried to minimize the potential negative consequences of omitting unobservable individual heterogeneity variables on the applied results. These two papers received significant extensions. The attention paid to Mundlak and to Balestra and Nerlove's works stemmed from the fact that they focused on modeling the error term and providing an economic meaning to the error term (as in the 1930s and the beginning of the 1940s), introducing then a break with the current interest and practice of the Cowles Commission, which was more focused on general statistical modeling. By not considering the error term as a nuisance, both papers raised the question of the possibility of getting information on the nature of the economic phenomena at stake. Thus they gave some revival

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

of the early debates within the econometric community on identification and estimation procedures, and more specifically the treatment of unobservable determinant and latent variables. Their innovative aspects are mainly the theoretical investigation of the error term based on economic concern, the recourse to random- or fixed-effects models, and the enlargement of the size of the sample. The general ANOVA model leads to the error component model. This model gave to econometricians the notion of specific effects as a key concept for linear panel data models.

Many articles have extended the work of Mundlak (1961) and Balestra and Nerlove (1966), including Mundlak 1963 and 1964, and Mundlak and Hoch 1965 in the similar spirit to that of Mundlak 1961. Later, Mundlak (1978a, 1978b) advocated for the fixed-effects approach, especially in his famous paper published in *Econometrica* (1978b). In 1981 Stephen Nickell demonstrated analytically the inconsistency of OLS in the coefficient of the lagged dependent variable in the fixed-effects regression, as for those of explanatory variables. On the other side, Nerlove (1967, 1971) verified the anomalies of Balestra and Nerlove 1966 using Monte Carlo simulations. Nerlove's results confirmed the inadequacy of several estimation techniques including OLS and ML. G. S. Maddala ([1975] 1994) underscored the importance of initial observations using ML on panel data. In 1978 Alain Trognon used the powerful results of Bernt Stigum (1974, 1976) to deduce the asymptotic properties of these estimators under various circumstances. The prevalence of boundary solutions for the ML appeared to be limited to some specific cases. Nevertheless, the main developments in understanding the role of initial conditions were provided by Theodore Wilbur Anderson and Cheng Hsiao (1981, 1982), followed by a thesis by Sevestre (1983), and articles by Alok Bhargava and John Denis Sargan (1983) and Sevestre and Trognon (1983, 1985).

These key references are strongly related to the emergence of panel data econometrics as a specific branch of econometrics. In the late seventies and eighties, French econometricians strongly influenced the promotion of panel data econometrics. In August 1977 Pascal Mazodier with Jacques Mairesse and Trognon were the first to organize a conference on panel data econometrics in Paris,[14] at INSEE under the auspices of CNRS.

14. See the special issue of *Annales de l'INSEE* titled "The Econometrics of Panel Data" (nos. 30–31, 1978). This volume contains twenty-five of the twenty-seven articles presented at the conference. Those articles were written by prestigious participants and some founding fathers of panel data econometrics: Gary Chamberlain, Robert Eisner, Zvi Griliches, Jerry A. Hausman, James J. Heckman, G. S. Maddala, Yair Mundlak, and Marc Nerlove, among others. Nerlove's introduction includes a selective summary of the conference papers.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

Ten years later, Sevestre and the members of ERUDITE started a conference series on the same topic with the support of Balestra, Badi Baltagi, Jean-Pierre Florens, Hsiao, Mairesse, and Trognon among others.[15] Since that time, this conference has become an important meeting place for presenting and discussing major research developments in panel data econometrics. The sixteenth edition organized by Maurice Bun, Jan Kiviet, and Tom Wansbeek took place in July 2010 at the University of Amsterdam. Since Mundlak's (1961) seminal paper, panel data econometrics has bloomed, and other topics related to this development need to be further investigated.

References

Airy, G. B. 1861. *On the Algebraical and Numerical Theory of Errors of Observations and the Combination of Observations*. Cambridge: Macmillan.

Anderson, T. W., and C. Hsiao. 1981. Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association* 76:598–606.

———. 1982. Formulation and Estimation of Dynamic Models Using Panel Data. *Journal of Econometrics* 18:47–82.

Balestra, P. 1967. *The Demand for Natural Gas in the United States: A Dynamic Approach for the Residential and Commercial Market*. Amsterdam: North-Holland.

Balestra, P., and M. Nerlove. 1966. Pooling Cross-Section and Time-Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas. *Econometrica* 34:585–612.

———. 1992. Formulation and Estimation of Econometric Models for Panel Data. In *The Econometrics of Panel Data: Handbook of Theory and Applications*, edited by L. Mátyás and P. Sevestre. Dordrecht: Kluwer Academic Publishers.

Bhargava, A., and J. D. Sargan. 1983. Estimating Dynamic Random-Effects Models from Panel Data Covering Short Time Periods. *Econometrica* 51:1635–59.

Chauvenet, W. 1863. *A Manual of Spherical and Practical Astronomy*. Vol. 1, *Spherical Astronomy*; vol. 2, *Theory and Use of Astronomical Instruments: Method of Least Squares*. Philadelphia: J B. Lippincott.

Daniels, H. E. 1939. The Estimation of Components of Variance. *Supplement to the Journal of the Royal Statistical Society* 6:186–97.

Eisenhart, C. 1947. The Assumptions Underlying the Analysis of Variance. *Biometrics* 3:1–21.

Epstein, R. J. 1987. A *History of Econometrics*. Amsterdam: North-Holland.

Fisher, R. A. 1918. The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh* 52:399–433.

15. Équipe de Recherche sur l'Utilisation des Données Individuelles Temporelles en Économie, University of Paris XII–Val de Marne.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

———. 1925a. *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.

———. 1925b. Theory of Statistical Estimation. *Proceedings of the Cambridge Philosophical Society* 22:700–725.

Frisch, R. A. K. 1933. Editorial. *Econometrica* 1:1–4.

———. [1938] 1995. Statistical versus Theoretical Relations in Economic Macrodynamics. In *Foundations of Modern Econometrics: The Selected Essays of Ragnar Frisch*, edited by O. Bjerkholt. Vol. 1. Aldershot, U.K.: Elgar.

Griliches, Z. 1957. Specification Bias in Estimates of Production Functions. *Journal of Farm Economics* 39:8–20.

Haavelmo, T. 1943. The Statistical Implications of a System of Simultaneous Equations. *Econometrica* 25:13–18.

———. 1944. The Probability Approach in Econometrics. *Econometrica* 12 (supplement): iii–vi, 1–115.

Halperin, M. 1951. Normal Regression Theory in the Presence of Intra-class Correlation. *Annals of Mathematical Statistics* 22:573–80.

Henderson, C. R. 1971. Comment on the Use of Error Components Models in Combining Cross Section with Time Series Data. *Econometrica* 39:397–401.

Hildreth, C. 1949. Preliminary Considerations Regarding Time Series and/or Cross Section Studies. Cowles Commission Discussion Paper: Statistics No. 333.

———. 1950. Combining Cross Section Data and Time Series. Cowles Commission Discussion Paper: Statistics No. 347.

Hoch, I. 1955. Estimation of Production Function Parameters and Testing for Efficiency. *Econometrica* 23:325–26.

———. 1957. Estimation of Agricultural Resource Productivities Combining Time Series and Cross-Section Data. PhD diss., University of Chicago.

———. 1958. Simultaneous Equations Bias in the Context of Cobb-Douglas Production Function. *Econometrica* 26:566–78.

———. 1962. Estimation of Production Function Parameters Combining Time Series and Cross-Section Data. *Econometrica* 30:34–53.

Koopmans, T. C. 1949. Identification Problems in Economic Model Construction. *Econometrica* 17:125–44.

———. 1950. *Statistical Inference in Dynamic Economic Models*. Cowles Commission Monograph No. 10. New York: Wiley.

Maddala, G. S. [1975] 1994. Some Problems Arising in Pooling Cross-Section and Time-Series Data. In *Econometric Methods and Applications*. Vol. 1. Aldershot, U.K.: Elgar.

Marschak, J., and W. H. Andrews. 1944. Random Simultaneous Equations and the Theory of Production. *Econometrica* 12:143–205.

Morgan, M. S. 1990. *The History of Econometric Ideas*. Cambridge: Cambridge University Press.

Mundlak, Y. 1961. Empirical Production Function Free of Management Bias. *Journal of Farm Economics* 43:44–56.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

———. 1963. Estimation of Production and Behavioral Functions from a Combination of Cross-Section and Time-Series Data. In *Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld*, edited by C. Christ et al. Stanford: Stanford University Press.

———. 1964. An Economic Analysis of Established Family Farms in Israel, 1953–1958. The Falk Project for Economic Research in Israel, Jerusalem.

———. 1978a. Models with Variable Coefficients: Integration and Extension. *Annales de l'INSEE* 30–31:483–509.

———. 1978b. On the Pooling of Time Series and Cross Section Data. *Econometrica* 46:69–85.

Mundlak, Y., and I. Hoch. 1965. Consequences of Alternatives Specifications in Estimation of Cobb-Douglas Production Functions. *Econometrica* 33:814–28.

Nerlove, M. 1967. Experimental Evidence on the Estimation of Dynamic Economic Relations from a Time-Series of Cross Sections. *Economic Studies Quarterly* 18:42–74.

———. 1971. Further Evidence on the Estimation of Dynamic Economic Relations from a Time-Series of Cross Sections. *Econometrica* 39:359–82.

———. 2002. *Essays in Panel Data Econometrics*. Cambridge: Cambridge University Press.

Nerlove, M., and P. Balestra. 1996. Formulation and Estimation of Econometric Models for Panel Data. In *The Econometrics of Panel Data: Handbook of Theory and Applications*, edited by L. Mátyás and P. Sevestre. 2nd ed. Dordrecht: Kluwer Academic.

Nickell, S. 1981. Biases in Dynamic Model with Fixed Effects. *Econometrica* 49:1417–26.

Qin, D. 1989. Formalization and Identification Theory. *Oxford Economic Papers* 41:73–93.

Qin, D., and C. L. Gilbert. 2001. The Error Term in the History of Time Series Econometrics. *Economic Theory* 17:424–50.

Scheffé, H. 1956. Alternative Models for the Analysis of Variance. *Annals of Mathematical Statistics* 27:251–71.

———. 1959. *The Analysis of Variance*. New York: Wiley.

Sevestre, P. 1983. Modèles dynamiques à erreurs composées. Doctoral thesis, Université Paris I–Panthéon Sorbonne.

Sevestre, P., and A. Trognon. 1983. Propriétés des grands échantillons d'une classe d'estimateurs des modèles autorégressifs à erreurs composées. *Annales de l'INSEE* 50:25–49.

———. 1985. A Note on Autoregressive Error Components Models. *Journal of Econometrics* 28:231–54.

Stigum, B. P. 1974. Asymptotic Properties of Dynamic Stochastic Parameter Estimates—III. *Journal of Multivariate Analysis* 4:351–81.

———. 1976. Least Squares and Stochastic Difference Equations. *Journal of Econometrics* 4:349–70.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023

Telser, L. G. 1964. Iterative Simultaneous Estimation of Sets of Linear Regressions. *Journal of the American Statistical Association* 59:845–62.

Trognon, A. 1978. Miscellaneous Asymptotic Properties of Ordinary Least Squares and Maximum Likelihood Methods in Dynamic Error Components Models. *Annales de l'INSEE* 30–31:631–57.

———. 2000. Panel Data Econometrics: A Successful Past and a Promising Future. Paper presented at the Ninth International Conference on Panel Data, Geneva, Switzerland, 22–23 June.

Wallace, T. D., and A. Hussain. 1969. The Use of Error Components Models in Combining Cross Section with Time Series Data. *Econometrica* 37:55–72.

Zellner, A. 1962. An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association* 57:348–68.

Downloaded from http://read.dukeupress.edu/hope/article-pdf/43/suppl_1/258/429941/HOPE435_12D-Kieffer_Fpp.pdf by ariane.dupont-kieffer@univ-paris1.fr on 21 May 2023